RHIC Computing Facility

DRAFT

Eric Lançon
### *August 23, 2016*

Outline

- Status of RCF, synergies with ATLAS Tier-1
- Performance in recent RHIC runs
- Future technological and data challenges
- Synergies with BNL Computing Initiative
- B725 infrastructure project

Status of RCF, synergies with ATLAS Tier-1

- RCF performed well during 2016 run
- Resources are ~fully utilised
- Hardware (CPU) is getting old, migration to new tape generation needed (space in HPSS)

- Synergy with ATLAS Tier-1
    - Economy of scale (operation, purchase,…)
    - Common procedures and configurations (resilience)
    - Common tools (batch system, storage, network)
    - Expertise from RCF benefit to ATLAS (and vis versa)
    - Access to Grid and cloud computing expertise
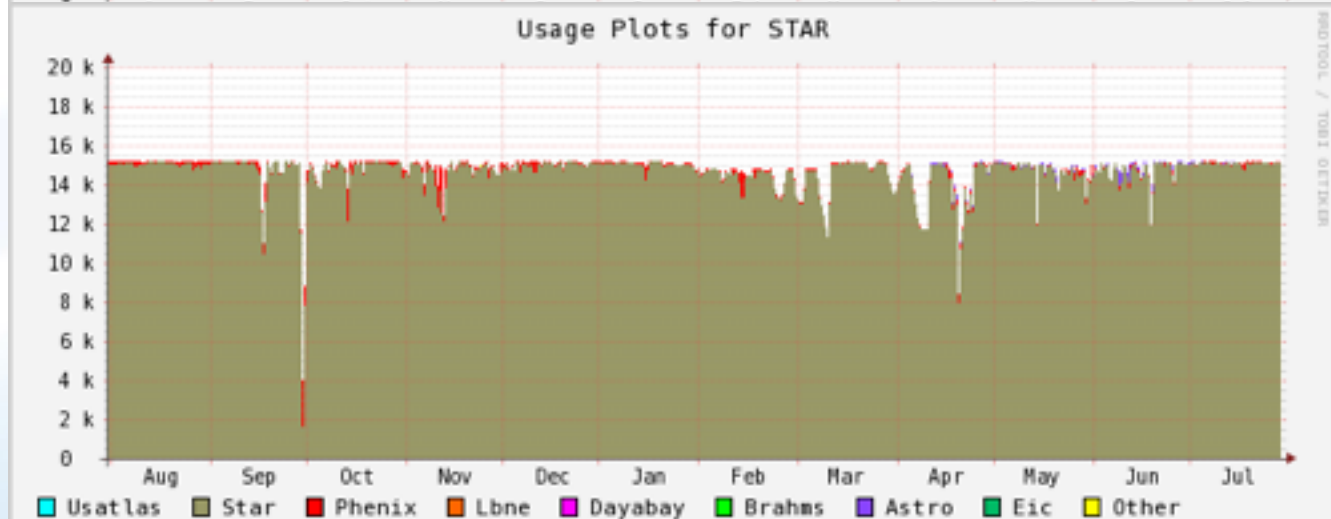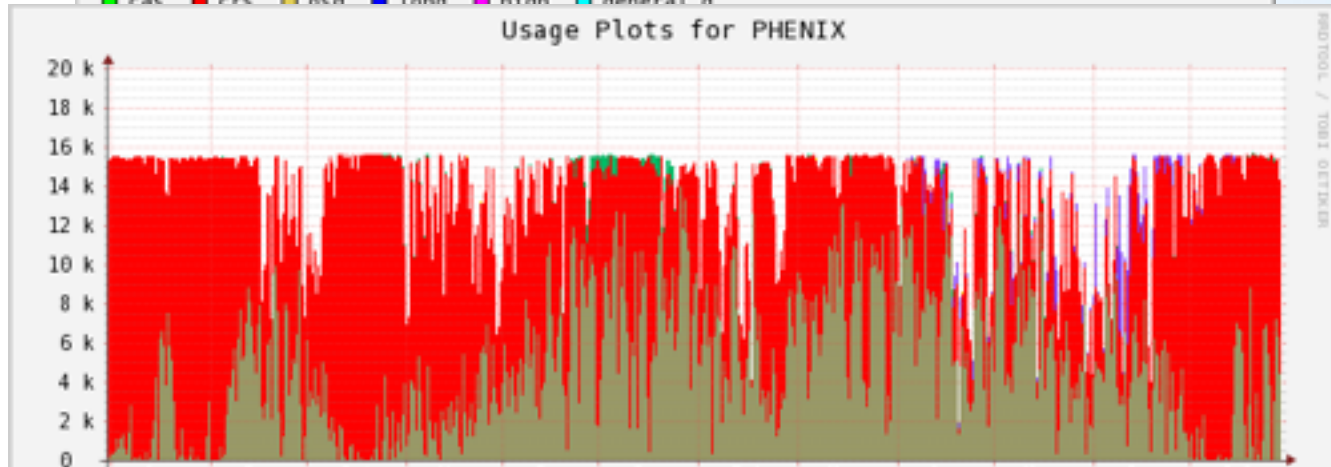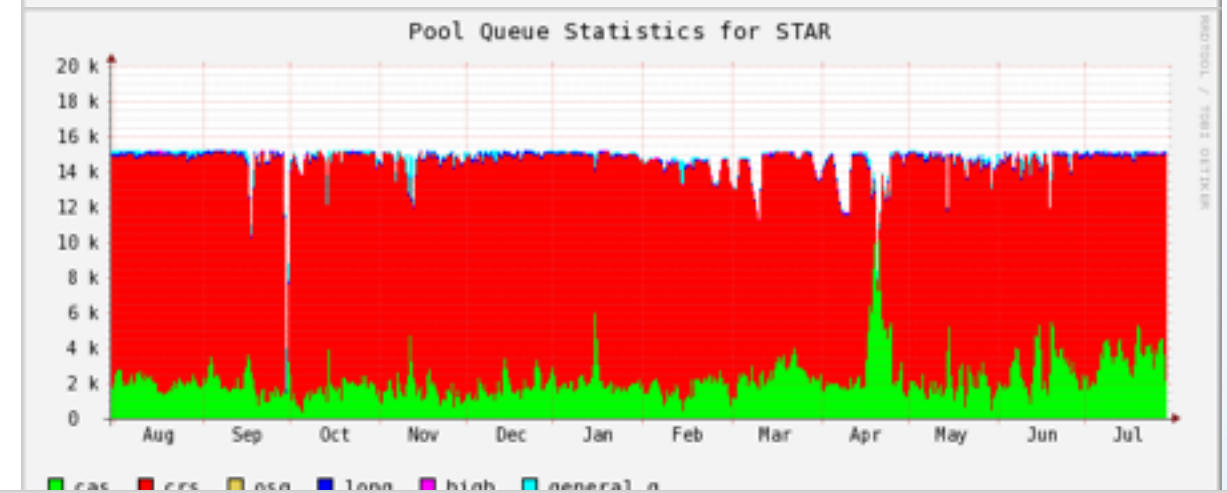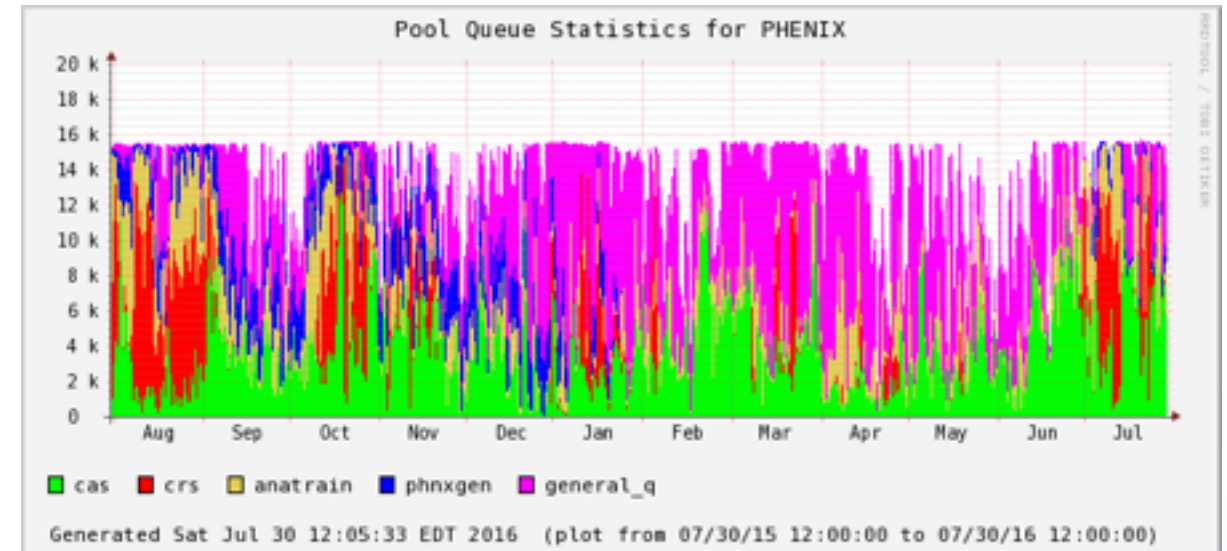
BROOKHAVEN
NATIONAL LABORATORY

Performance in recent RHIC runs

- No issue in
    - data transfer from experiments to facility
    - Writing to tape
- give some numbers
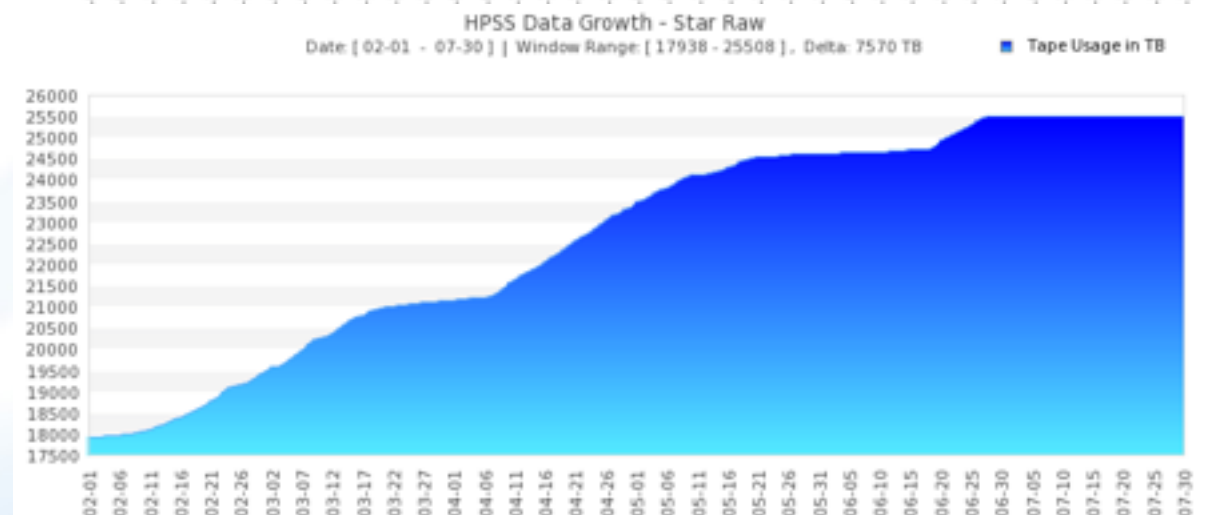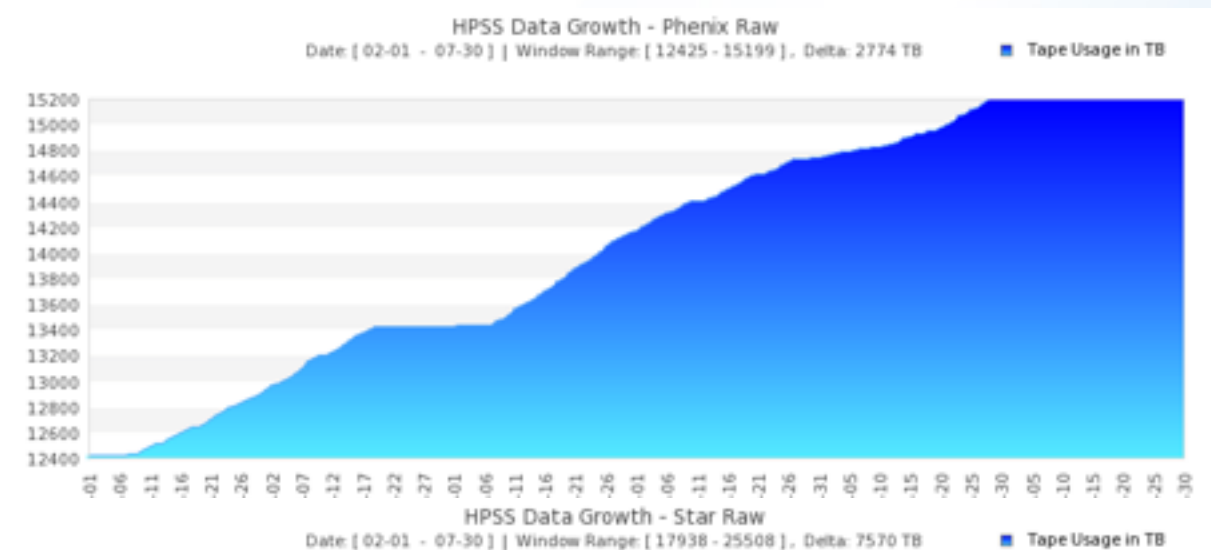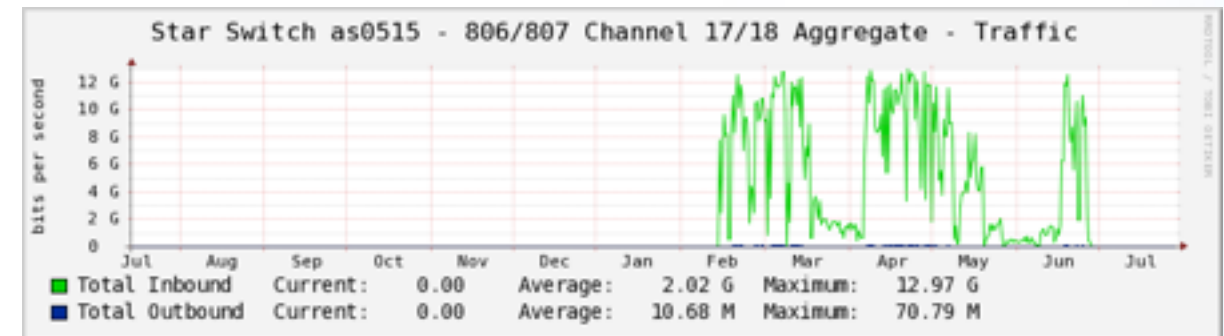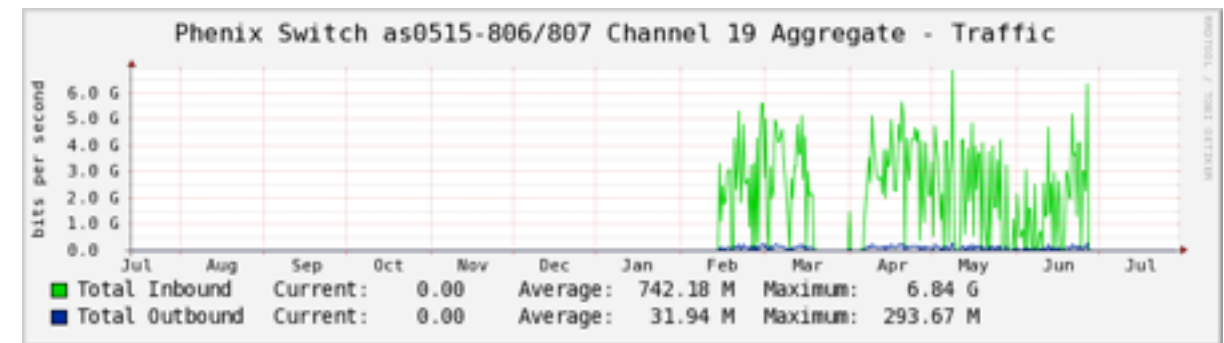
CPU usage of the farms

- STAR farm almost continuously saturated while PHENIX farm is not
- PHENIX farm used by STAR analysis jobs
  - Optimisation of batch system (Condor) performed, STAR workflow optimisation to be done
  - STAR reconstruction cannot use PHENIX resources because STAR storage is distributed on CPU nodes
- Lessons for the future
  - Computing models (workflow management, data organisation, …) and technological choices (storage, CPU,…) of experiments should not be too different in order to benefit from a global pool of resources

Pool Queue Statistics for PHENIX

cas crs anatrain phnxgen general_q

Generated Sat Jul 30 12:05:33 EDT 2016 (plot from 07/30/15 12:00:00 to 07/30/16 12:00:00)

Pool Queue Statistics for STAR

Usage Plots for PHENIX

Usage Plots for STAR

Usatlas Star Phenix Lbne Dayabay Brahms Astro Eic Other

Future technological and data challenges

- Future of computing is multi-core
  - New hardware are multi-core 8, 16, 32, 64,…. with less memory per core (or at a higher price)
  - Could software of RHIC experiments be multi-core? is it worst for existing experiments?
- Object store technology, ATLAS is migrating
- RHIC hardware is getting old ~22% older than 6 years
- Tape technology
  - 2 generations behind in tape technology
  - Move to new generation (LT07) to gain in capacity / media
  - ONLY one copy of RAW on tape
    - In moving data to new media 2 copies are made
    - xxx months for migration
    - But the 2 copies will be in the same building

Computational Science Initiative : CSI

- **CSI** *: Integrating data-intensive science expertise and investments across the Laboratory to tackle "big data" challenges*
  - Leverage investments across multiple programs
  - Patterns : universities (Columbia, Cornell, New York University, Stony Brook, and Yale) and companies including IBM Research.
- Scientific Data and Computing Center : **SDCC**
- Computer…
- Computational…
- Center for….

Computational Science Initiative Directorate

	KERSTIN KLEESE VAN DAM* *Director*	
HR BUSINESS PARTNER *Christel Colon*	Robert Harrison *Chief Scientist*	**RESEARCH OPERATIONS** *Linda Bowerman*
BUSINESS OPERATIONS *Ken Koebel*	Lauri Peragine *Administrative Specialist*	**DIRECTORATE CHIEF OPERATIONS OFFICER** *Ann Emrick*

COMPUTER SCIENCE AND MATHEMATICS *BARBARA CHAPMAN* *Chair*	**BNL SCIENTIFIC DATA AND COMPUTING CENTER** *ERIC LANÇON* *Chair*	**COMPUTATIONAL SCIENCE LABORATORY** *NICK D' IMPERIO* *Chair*	**CENTER FOR DATA DRIVEN DISCOVERY (C3D)** *KERSTIN KLEESE VAN DAM* *Director*

*Reports to Robert Tribble, Deputy Director Science & Technology
ESH Coordinator – Bob Colichio, Pat Carr

Approved: Kerstin Kleese van Dam   Date: June 2, 2016

BROOKHAVEN
NATIONAL LABORATORY

Brookhaven Science Associates

Scientific Data and Computing Center : SDCC

- Within CSI : an integrated scientific data, computing and networking infrastructure across BNL
- RACF the RHIC and ATLAS Computing Facility is the core of SDCC :
  - A world class HEP and NP facility
  - Expertise and infrastructure necessary for leadership & innovation in data-driven discovery
- SDCC is operating
  - The Lab's new Institutional Computing system (IC)
  - A series of novel architecture research systems,
  - As well as computing and data services for other third-party clients.

BROOKHAVEN
NATIONAL LABORATORY

Synergies with BNL Computing Initiative

- CSI is purchasing or complementing purchases in the area of HPC computing
  - Institutional cluster (Fall 2016)
  - Knight Landings (KNL) Intel farm (Fall 2016). Initiated by BNL QCD group and RIKEN, CSI doubled the capacity.
- These resources will be made available to RHIC program in opportunistic mode
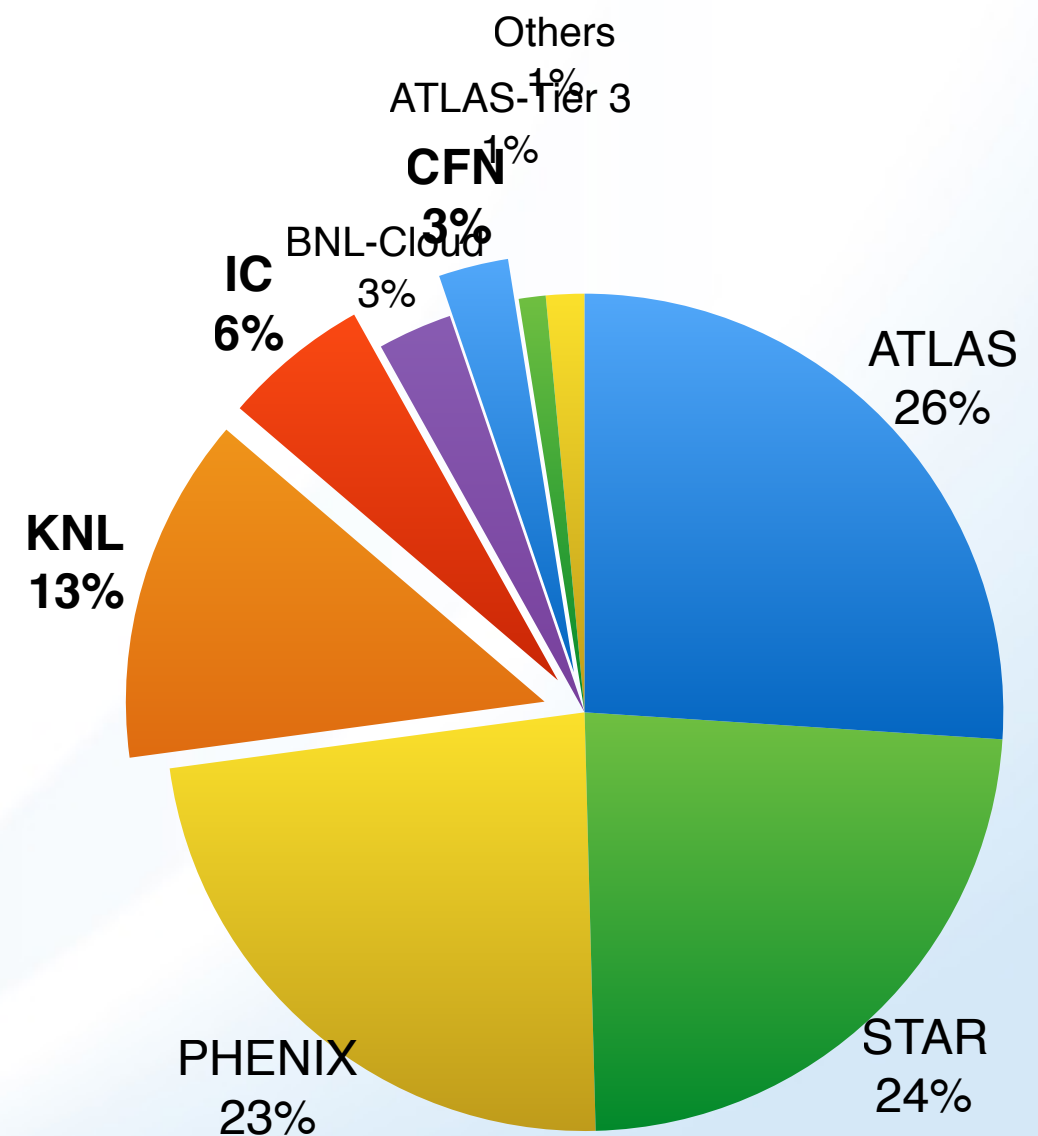- Issue : manpower to port RHIC codes?

SDCC evolution over the next months

- **Institutional Cluster (IC)**
  - Fall 2016 : 108 nodes (Intel Xeon + Nvidia K80 GPU + InfiniBand)
  - 2 x mid-2017
- **Intel Knights Landing (KNL) cluster :**
  - Fall 2016 : 144 nodes (Intel Xeon Phi + Omni-Path)
  - for RBRC (RIKEN and BNL Research Center)
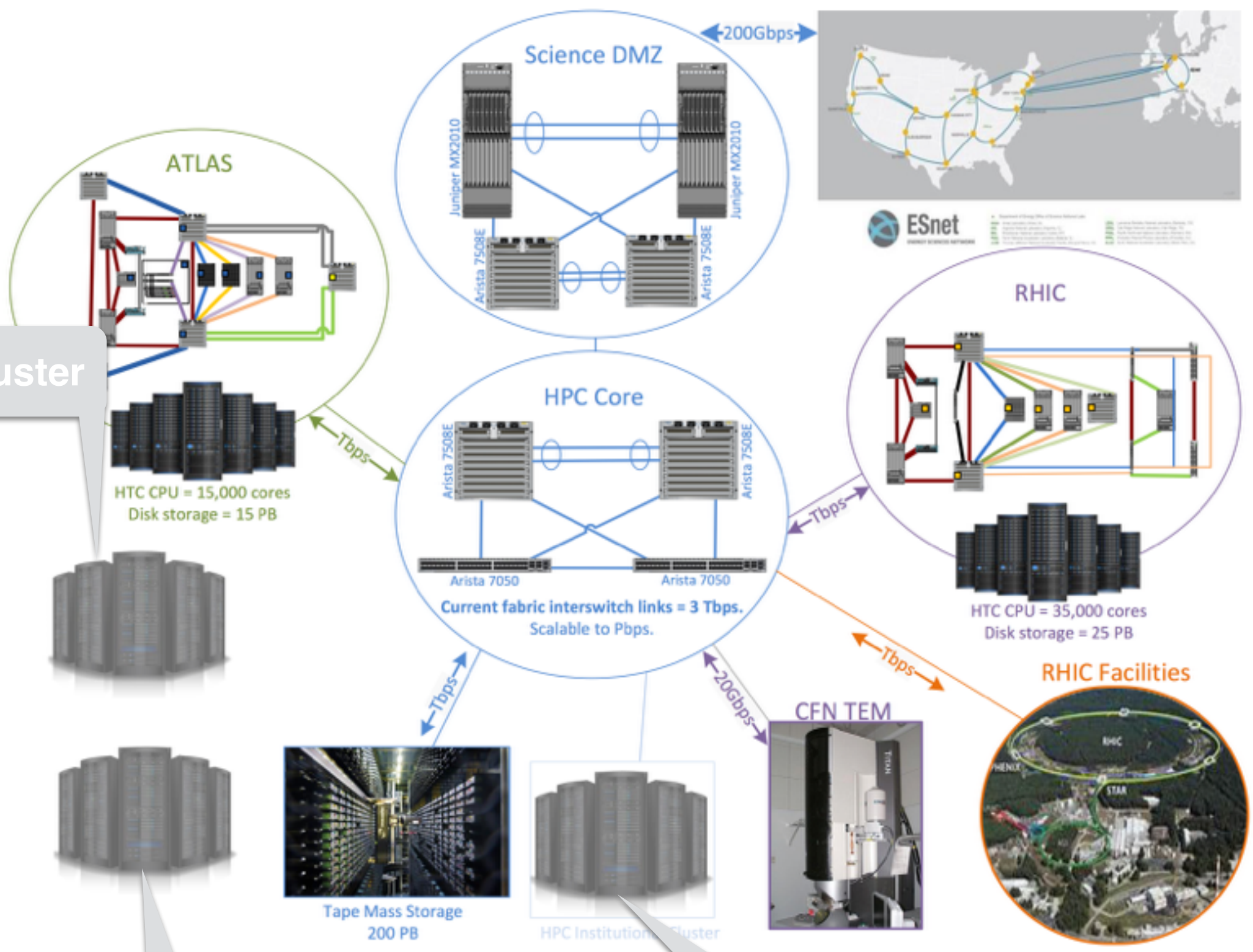 - and for CSI partners

End of 2016
70k cores
HPC : 22%

Pie chart:
- ATLAS 26%
- STAR 24%
- PHENIX 23%
- KNL 13%
- IC 6%
- BNL-Cloud 3%
- CFN 3%
- ATLAS-Tier 3 1%
- Others 1%

KNL cluster

ATLAS

Science DMZ

◄—200Gbps—►

Juniper MX2010

Juniper MX2010

Arista 7508E

Arista 7508E

ESnet
ENERGY SCIENCES NETWORK

RHIC

HTC CPU = 15,000 cores
Disk storage = 15 PB

◄—Tbps—►

HPC Core

Arista 7508E

Arista 7508E

◄—Tbps—►

Arista 7050

Arista 7050

Current fabric interswitch links = 3 Tbps.
Scalable to Pbps.

HTC CPU = 35,000 cores
Disk storage = 25 PB

RHIC Facilities

◄—Tbps—►

◄—20Gbps—►

◄—Tbps—►

CFN TEM

Tape Mass Storage
200 PB

HPC Institutional Cluster

USQCD cluster

Institutional Cluster

Brookhaven Science Associat

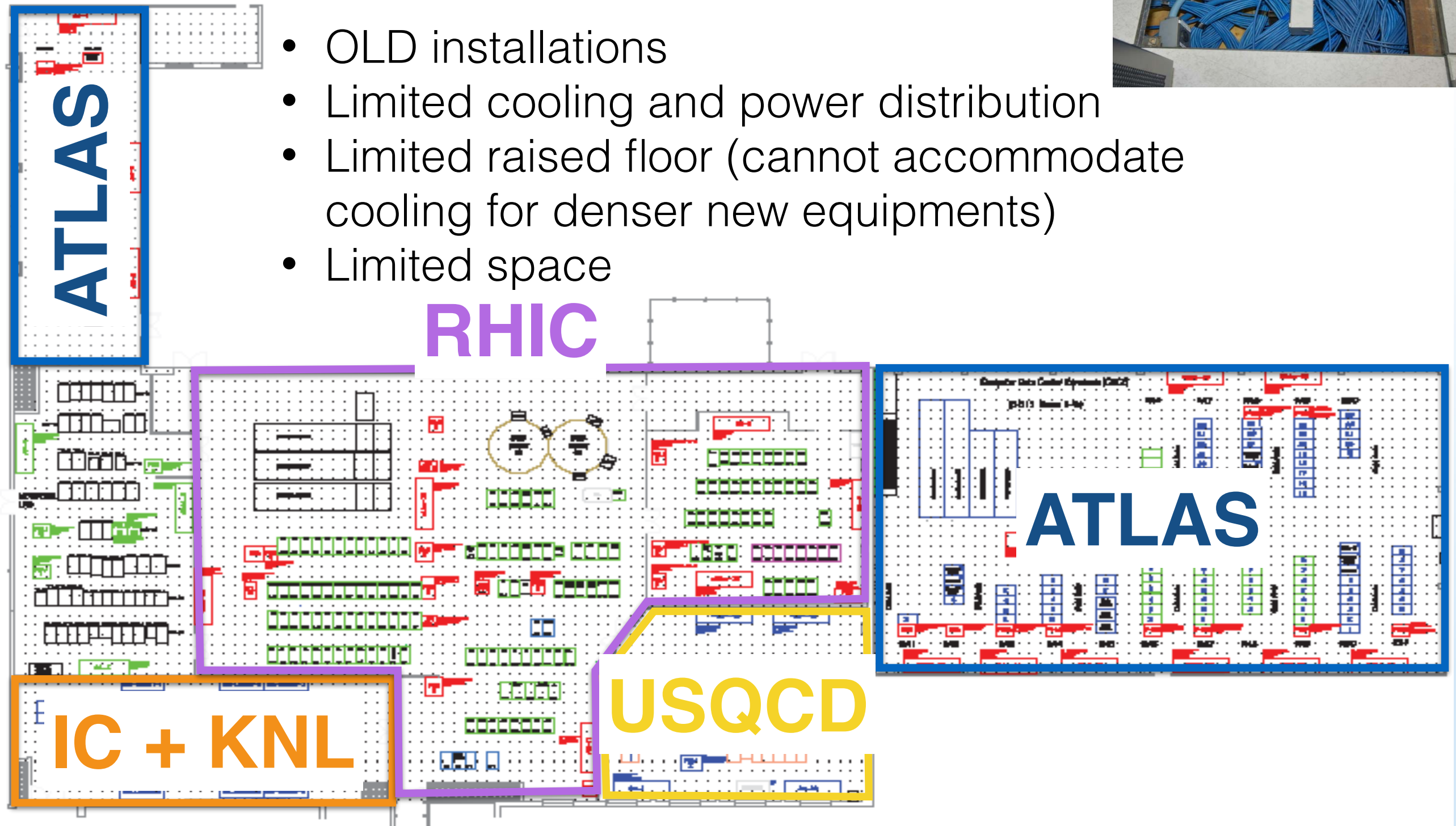BROOKHAVEN
NATIONAL LABORATORY

B725 infrastructure project

Computing room(s) - Issues

- OLD installations
- Limited cooling and power distribution
- Limited raised floor (cannot accommodate cooling for denser new equipments)
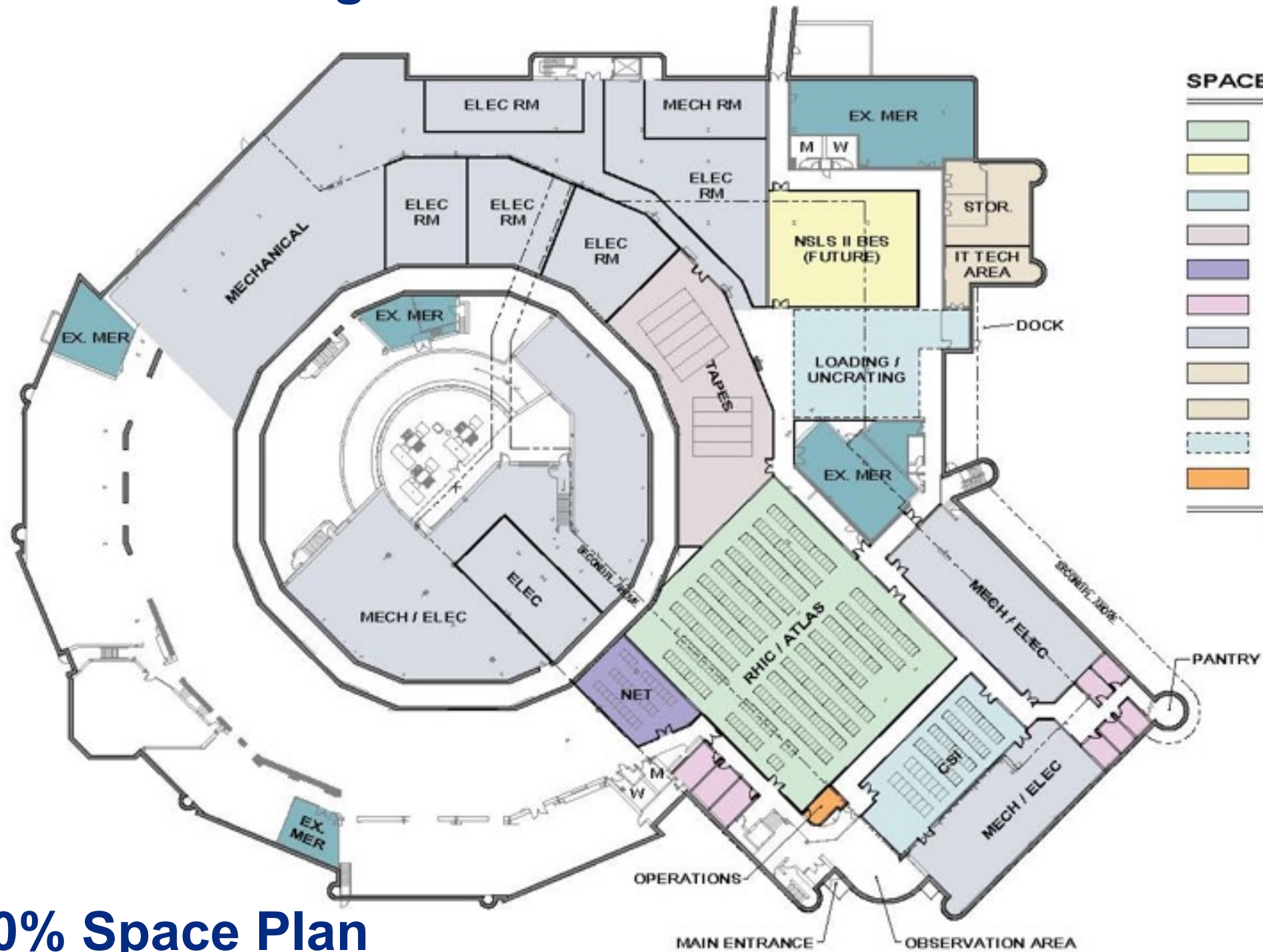- Limited space

RHIC

ATLAS

ATLAS

IC + KNL

USQCD

New computing room

Core Facility Revitalisation – Conceptual Design

NSLS-I Building

60% Space Plan

CD-1 review end of August

BROOKHAVEN
NATIONAL LABORATORY

CFR – Preliminary Schedule

Preliminary CFR Funding Analysis - 1 Yr CR 2017 (Renovation Alternative)

Facility migration

Proposed run schedule for RHIC
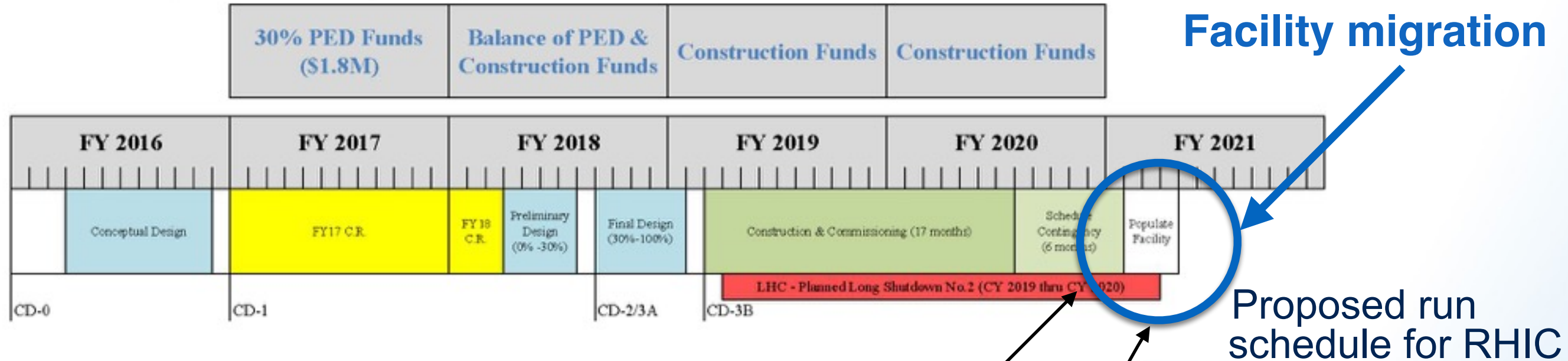
Longer term LHC schedule

The outline LHC schedule out to 2035 presented by Frederick Bordry to the SPC and FC June 2015 can be found here
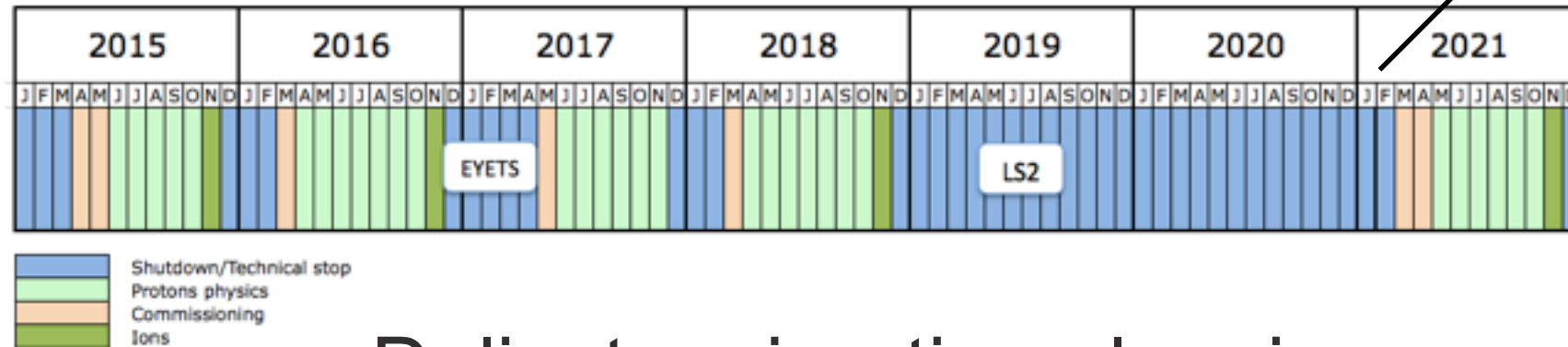
- Delicate migration planning

- In phase with LHC Long Shutdown 2 and RHIC

  - ATLAS Tier 1 MoU : Uptime 99%

Years	Beam Species and Energies
2016	High statistics Au+Au d+Au energy scan
2017	High statistics Pol. p+p at 510 GeV
2018	$^{96}Zr+^{96}Ru$ isobar run
2019-20	7.7-20 GeV Au+Au (BES-2)
2021	No Run ?
2022-23	200 GeV Au+Au with upgraded detectors Pol. p+p, p+Au at 200 GeV
2024---	Program TBD

CFR Design – An Incremental Approach

- **Power**
  - Day-one capability (2021) – 2.4 MW IT power (dedicated computing power). This is approximately double current RACF IT power.
  - Provide provision for future1.2 MW IT power increments to 6MW Max.

- **Cooling**
  - Day-one cooling capability to support 2.4 MW IT power
  - Provide provision for future 1.2 MW IT power deployments

- **Space**
  - Day-one - Accommodate approximately 33% footprint expansion (Racks) within defined spaces.
  - Day-one - Accommodate approximately 3,500 SF additional, unassigned space.
  - Provide opportunity for future (long term) growth within the balance of the 725 facility. Both computing and offices.

Summary